Sustained Spread of HIV-1 Drug Resistance in Treatment-Naïve Patients in the United Kingdom.

Jean L. Mbisa[1], Esther Fearnhill[2], David T. Dunn[2], Deenan Pillay[3], David Asboe[4], Patricia A. Cane[1], on behalf of the UK HIV Drug Resistance Database.

[1]Public Health England, [2]MRC Clinical Trials Unit, [3]University College London Medical School, [4]Chelsea and Westminster Hospital, London, UK

INTRODUCTION

Transmitted drug resistance (TDR) in those infected with HIV-1 subtype B and tested between diagnosis and initiation of ART peaked in 2002 at 15.5% in the UK. It then declined to a nadir of ~9% in 2007 and has since remained relatively stable [1]. The presence of TDR has been shown to be associated with a higher risk of failure of first line therapies, presumed to be due to these therapies being rendered sub-optimal by the presence of drug resistance mutations [2]. However, many of the mutations found in TDR are thought not to compromise current first line therapies with the exception of non-nucleoside reverse transcriptase inhibitors, due to the move away from zidovudine- and stavudine-based therapies [1].

The stability of TDR mutations in untreated patients has been examined [3,4]. It has been found that while mutations such as K65R and M184V in reverse transcriptase are lost rapidly due to fitness costs, other mutations can persist for many years. T215Y/F mutations in reverse transcriptase, commonly observed in patients failing therapies including AZT or d4T, are quickly replaced with "revertant" mutations (T215rev) such as T215D/C/E/S, following transmission [5]. These T215rev mutations have been found to persist for a median of 13.0 years (95% confidence interval 6.6-25.7 years) [4]. In recent years, most (60-70%) resistance tests from treated patients in the UK show no resistance mutations and there is a marked discordance between those mutations seen in failing patients and those seen in TDR [6]. With the longevity of some TDR mutations, a substantial proportion of TDR may therefore be derived from untreated and possibly undiagnosed patients. Onward transmission of TDR was first reported in 2003 [7], and an analysis of the UK HIV resistance database in 2009 found five treatment-independent clusters of TDR [8]. In the Swiss HIV cohort study, it has now been shown that treatment-naïve individuals are the major source of TDR in men who have sex with men (MSM) [9].

This report focuses on the source of the most common TDR mutations seen associated with resistance to NRTIs, NNRTIs and protease inhibitors namely at codons T215, K103 and L90 which most recently have been reported in 4.01%, 2.79% and 0.95%, respectively, of all new subtype B diagnoses in the UK. It shows that a substantial proportion of TDR sequences with these most common mutations cluster together indicating considerable onward transmission.

METHODS

Patients. All treatment-naïve subtype B sequences with any of L90M, K103N, or T215 mutations (Y/F/revertants). Initially sequences from 1204 patients were selected. From these 45 were removed due to identical dates of birth and <1.5% genetic distance. In addition 9 were removed as considered due to laboratory error, leaving sequences from 1140 patients. Of these 78% were male and 70% MSM. The sample years were 1997-2011 with the majority (923) from between 2004 and 2010.

Phylogenetic analysis. A Fasttree maximum likelihood (ML) tree was created using the GTR + gamma model of nucleotide substitution, following removal of resistance positions. Clusters were selected with a maximum genetic distance of <= 1.5% and <= 2% using TN93 distance method with partial inclusion of mixed bases. 1.5% is <0.01 percentile of distribution of all pairwise distances for these sequences. Clusters with support of < 0.8 were analysed at higher nodes to rule out errors due to distance cut-off precision. The process was then repeated including 1000 randomly selected first naïve sequences from different subtype B patients to check that the clusters were sustained. Where clusters were found to include pairs of sequences with zero genetic distance, these were manually checked with particular reference to patient demographic data such as date of birth and any sequence likely to have come from the same patient was removed (9 in all).

Estimating the time of origin and basic reproduction rate (R_0) of TDR clusters. The time of most recent common ancestor (tMRCA) of the large transmission clusters with 8 or more sequences, which corresponds to the date at which resistance mutations occurred in the clusters, was estimated using a Bayesian Markov chain Monte Carlo (MCMC) approach as implemented in BEAST software. For computational reasons, the 624 sequences that formed highly supported clusters by ML were randomly divided into smaller groups of ~150 sequences keeping sequences in clusters together and this was repeated several times to test the robustness of the clusters. The analysis was preceded by a comparison of different molecular clock (strict or relaxed), demographic (BSP) and nucleotide substitution (GTR+gamma or SRD06) using Bayes Factor (with a BF >20 level of significance). R_0 estimates were carried out using the birth-death model (BDM) as implemented in BEAST [10]. All estimates were done by sampling every 5,000th generation from an MCMC run of 50,000,000 generations.

RESULTS

Table 1. Proportion of sequences harbouring the TDR mutations *PR* L90M, and *RT* K103N and T215Y/rev that were analyzed in this study. These sequences represent ~50% of all TDR sequences in the database

Mutation	No. of sequences (percentage of total)
L90M	109 (9.6%)
K103N	359 (31.5%)
T215rev	540 (47.4%)
T215Y	42 (3.7%)
L90M + K103N	12 (1.1%)
L90M + T215rev	19 (1.7%)
K103N + T215rev	47 (4.1%)
L90M + K103N + T215rev	12 (1.1%)
Total	**1140 (100%)**

Table 2. Distribution of mutations in the clusters using a maximum genetic distance cut-off of ≤ 1.5%. This shows that all 3 single mutations have at least one cluster with 10 or more sequences.

	Mutation(s)							
Cluster size	L90M	K103N	T215rev	L90M + K103N	L90M + T215rev	K103N + T215rev	L90M + K103N + T215rev	Total
2	9	26	64		1	1	1	**102**
3	3	15	22		1			**41**
4	1	4	11					**16**
5	1	6	6					**13**
6		2	1			1		**4**
7	1	1	2					**4**
8			2			1		**3**
9			2					**2**
10		2						**2**
11	1							**1**
12			1					**1**
15	1							**1**
Total	**17**	**56**	**111**	**0**	**2**	**3**	**1**	**190**

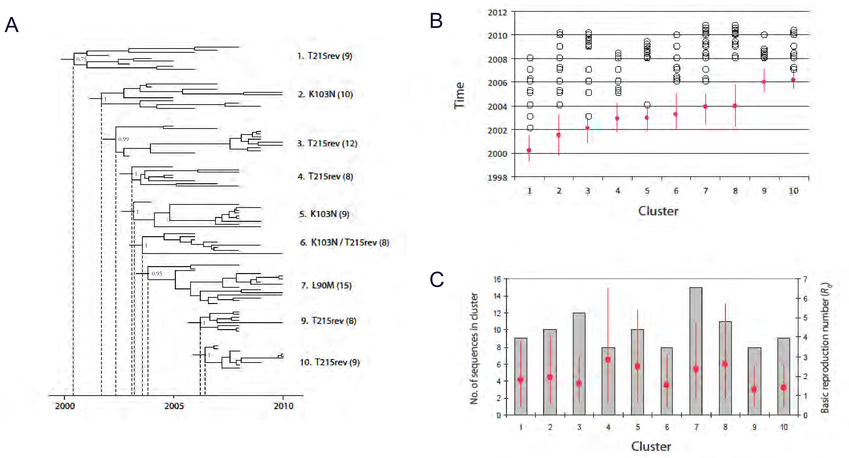

Figure 2. Estimating the time of origin and basic reproductive number (R_0) of large transmission clusters (≥ 8 sequences) using a Bayesian MCMC approach. (A) Phylogenies of large clusters showing the time of origin. The mutation(s) and the number of sequences (in brackets) contained in each cluster are indicated next to each cluster. The Bayesian posterior probabilities for the node subtending each cluster is shown next to the node. (B) Graphical representation of the time of origin of the clusters (red dot) with 95% highest probability density (HPD) interval (red line) and time of sampling of the data in the clusters (black circles). The time of origin of the clusters was estimated to range between 2000 and 2006. (C) Estimation of the R_0 i.e. the number of cases one case generates over the course of its infectious period using a Bayesian method of phylogenetic analysis based on the Birth-Death Model. An R_0 >1 indicates an ability of the infection to spread. The graph shows the R_0 for each cluster ranging from 1.3 to 2.8 (red dot) with 95% HPD intervals (red line). The gray bars indicate the number of sequences in each cluster. The clusters are numbered 1 to 10 in all panels. Cluster no. 8 (containing 11 sequences with L90M mutation) is the only large cluster that was not maintained when sequences were randomly divided into smaller groups of ~150 sequences for Bayesian phylogenetic inference, and therefore is not represented in panel A.

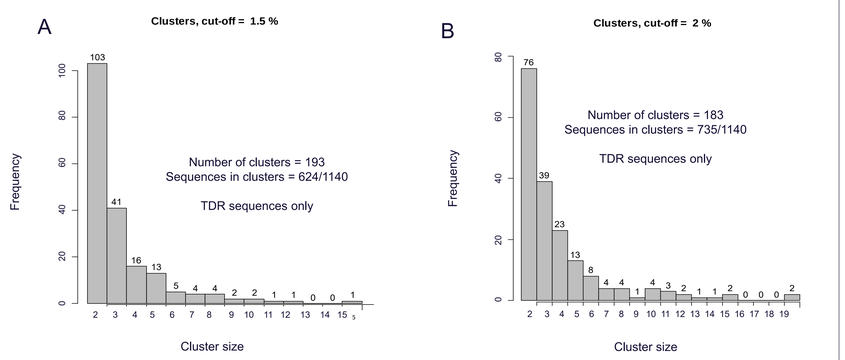

Figure 1. Effect of varying genetic distance threshold cut-off on number and size of clusters. Clusters were selected with a maximum genetic distance of ≤ 1.5% (A) and ≤ 2% (B) using TN93 distance method. Each figure shows the number of TDR sequences with L90M, K103N or T215rev mutations that fell into clusters and the distribution of cluster size. At 1.5% cut-off, 193 clusters were observed ranging in size from 2 to 15 sequences with 115 (61%) of the clusters containing more than 2 sequences. At 2% cut-off, 183 clusters were observed ranging in size from 2 to 19 sequences with 107 (59%) of the clusters containing more than 2 sequences

CONCLUSIONS

A considerable proportion of TDR in the UK (with mutations PR L90M, and RT K103N and T215rev) is linked to other TDR . This is in agreement with a recent smaller study of the Swiss cohort [9]. Thus a high proportion of TDR is probably derived by onward transmission from treatment-naïve patients. This is not unexpected since 82% of new infections in MSM are estimated to be derived from undiagnosed (and therefore untreated) individuals [11] while the proportion of patients failing treatment with resistance is small (~30%) [6].

A decrease in the prevalence of TDR could be achieved by enhanced testing to increase early diagnosis and to reduce onward transmission.

ACKNOWLEDGEMENTS

We acknowledge the UK HIV Drug Resistance Database Steering Committee and Contributing Centres (listed on www.hivrdb.org.uk)

REFERENCES

1. UK Collaborative Group on HIV Drug Resistance BMJ 2012, **345**:e5253
2. Wittkop *et al*; Lancet Infec Dis 2011; **11**:363
3. Pao *et al*; J Acquir Imune Def Syndr 2004; **37**:1570
4. Castro *et al*; JID 2013; **208**:1459
5. Yerly *et al*; J Virol 1998; **72**:3520
6. UK Drug Resistance database; http://www.ctu.mrc.ac.uk/hivrdb/public/surveillance.asp
7. Taylor *et al*; AIDS Res Human Retroviruses 2003; **19**:353
8. Hue *et al*; J Virol 2009; **83**:2645
9. Drescher *et al*; CID 2014; **58**:285
10. Stadler *et al*; Mol. Biol. Evol. 2012; **29**:347
11. Phillips *et al*; Plos One 2013; **8**:e55312